植 物 分 类 与 资 源 学 报　2013，35 (6)：693～700
Plant Diversity and Resources　　DOI：10.7677/ynzwyj201313186

应用 DNA 条形码对植物标本的核查*

曾春霞，贺正山，杨俊波**
（中国科学院昆明植物研究所中国西南野生生物种质资源库，云南 昆明　650201）

摘要：DNA 条形码主要目的是物种鉴定和新物种或隐存种的发现，而 DNA 条形码参考数据库是物种快速鉴定的重要基础。目前中国维管植物 DNA 条形码参考数据库正在建设之中，借助于公共数据库（NCBI）和初步建立的中国植物 DNA 条形码参考数据库，运用 DNA 条形码数据开展了植物标本鉴定的核查工作：（1）比较 DNA 序列信息与标本鉴定信息，从科、属、种级水平查找鉴定错误的标本；（2）基于有较好研究基础的 DNA 条形码参考数据库，开展未知标本的鉴定；（3）通过对标本核查的总结，提出 DNA 条形码参考数据库建设过程中的几点建议。

关键词：植物；标本鉴定；DNA 条形码；参考数据库
中图分类号：Q 949，Q 78.1　　**文献标识码**：A　　**文章编号**：2095-0845(2013)06-693-08

DNA Barcoding for Plant Specimen Identification*

ZENG Chun-Xia，HE Zheng-Shan，YANG Jun-Bo**
(*Germplasm Bank of Wild Species in Southwest China*，*Kunming Institute of Botany*，*Chinese Academy of Sciences*，Kunming 650201，China)

Abstract：DNA barcoding comprises two distinct aims：specimen identification and species discovery. To achieve these goals，a DNA barcode reference library is indispensable and is currently under construction for Chinese plants. In this paper，we used the existing public database，NCBI，and also our DNA barcode reference library，which is under construction，to identify plant specimens. By combining DNA barcode sequences with scientific names determined by taxonomists，we first attempted to find incorrectly identified specimens and also to check specimens whose DNA sequences were wrongly labeled. Secondly，we attempted the identification of unknown specimens in some groups for which DNA barcodes are available. Thirdly，we make some suggestions for the development of a plant DNA barcode reference library.

Key words：Plants；Specimen identification；DNA barcoding；Reference library

对物种的准确鉴定是分类学的主要目的，也是物种多样性管理、保护和可持续利用的迫切需求（Chase 和 Fay，2009）。自林奈的双名法确立以来，已鉴定和描述了约 170 万种生物物种（Hawksworth，1995），但仍有 86% 的陆地生物和 91% 的海洋生物有待发现和描述（Mora 等，2011）。目前，传统分类学对脊椎动物、昆虫及高等植物已开展了广泛研究，然而，对于形态特征不明显的微小生物研究较少（任保青和陈之端，2010）。随着社会经济的发展，人们对生物多样性的认识、保护和利用越来越关注，对物种快速准确鉴定的需求也日益迫切（Che 等，2010）。然而传统分类学家的数量正在不断缩减，因此，对地球上如此众多物种的认识和准确鉴定是当前面临的艰巨任务之一（Blaxter，2003；Tautz 等，2003；Li 等，2011b）。

DNA 条形码技术是利用标准化的 DNA 序列

* 基金项目：科技部科技基础性工作专项项目（2013FY112600）；国家高技术研究发展计划（863 计划）主题项目（2012AA021801）；中国科学院仪器设备功能开发技术创新项目实施方案项目（植物遗传信息高效获取技术体系集成研制）
** 通讯作者：Author for correspondence；E-mail：jbyang@mail.kib.ac.cn
收稿日期：2013-09-11，2013-10-16 接受发表
作者简介：曾春霞（1980-）女，博士，主要从事植物系统发育研究。E-mail：zengcx@mail.kib.ac.cn

对物种进行快速准确的鉴定，其实质是作为一种新的性状来构建分类系统，实现序列本身变异信息与现有形态分类学的结合（DeSalle等，2005；Kristiansen等，2005；Will等，2005；Haase等，2007；Hajibabaei等，2007；Vogler和Monaghan，2007）。该技术不完全依赖形态特征保存完整的标本，对于已经失去了重要形态分类特征的待鉴定标本，也能实现正确鉴定，它是对传统形态分类学强有力的补充（Godfray，2002；Mallet和Willmott，2003；Hebert和Gregory，2005；Schindel和Miller，2005）。同时，DNA条形码不仅可以鉴定物种、辅助分类学修订，而且有利于新种和隐存种的发现，为生物多样性的研究提供了新的研究手段和思路（Dayrat，2005；Fitzhugh，2006；Agnarsson和Kuntner，2007）。

自Hebert等（2003）提出DNA条形码用于物种快速鉴定以来，植物DNA条形码研究也取得了长足发展。2009年，国际生命条形码联盟植物工作组（CBOL Plant Working Group，2009）推荐*rbc*L+*mat*K组合作为陆地植物的核心DNA条形码，用于构建植物物种鉴定的统一框图。第三界国际生命条形码大会上，ITS和*trn*H-*psb*A被建议为植物条形码的辅助条码（Hollingsworth等，2011）。中国植物条形码研究组（China Plant BOL Group）基于中国种子植物的75科141属1 757种约6 286个样品，通过对4个条码（*rbc*L、*mat*K、*trn*H-*psb*A和ITS）综合比较分析，建议ITS（或ITS2）可作为种子植物的核心条码之一（Li等，2011a）。DNA条形码数据分析采用3种常用的方法，即基于BLAST分析的相似法（Meier等，2006）、距离法（distance）和基于遗传分化距离的建树法（tree-building）（Matz和Nielsen，2005；Nielsen和Matz，2006）。

为了实现对物种的快速鉴定，DNA条形码参考数据库（Reference Library）的构建是物种鉴定的重要前提。目前分子鉴定和系统学分析中常用的DNA数据库有：美国国家生物技术信息中心（National Center for Biotechnology Information，GenBank），欧洲分子生物学实验室（European Molecular Biology Laboratory，EMBL）和日本核酸数据库（DNA Data Bank of Japan（DDBJ)。为了进一步推进DNA条形码研究，生命条形码联盟（CBOL，Consortium for the Barcode of Life）建立了生命DNA条形码数据库（BOLD，The Barcode of Life Data），为专门收集和分析DNA条形码数据的平台。最新统计，约181 263种物种已经完成DNA条形码信息的描绘，同时具DNA条形码序列信息的标本也有2 239 378份（http://www.barcodinglife.com/index.php/TaxBrowser_Home），该数据库以鱼类、鸟类和昆虫类为主，真菌和植物DNA条形码信息相对较少。中国科学院昆明植物研究所联合全国相关科研院校开展了中国重要植物类群的采集、DNA条形码的测定与分析等工作，已经积累了中国维管束植物约230科1 600属6 700种约60 000条DNA条形码序列，目前DNA条形码参考数据库已初具雏形。

借助于公共数据库NCBI和初步建立的DNA条形码参考数据库，我们应用DNA条形码技术开展了植物标本物种鉴定的核查工作。

1 有鉴定信息标本的再确认

传统植物鉴定主要依据植物的形态特征，参考地理分布，以分类检索表为路径，根据植物志属进行物种匹配（方伟和刘恩德，2012）。由于需要有较强的专业背景，分类学家需要经过多年的培养才能擅长于某一门类。我国植物多样性丰富，即使是有经验的分类学家也难免鉴定错误。DNA条形码的工作流程涉及生物标本的采集、保存和规模化管理；样品DNA的提取；条形码标准基因的设计和筛选；目标片段扩增、测序；序列分析和数据提交到数据库等。利用样品已经获得的DNA条形码序列信息，基于BLAST和建树法，对DNA条形码序列和标本信息逐条比对核查，可有效查找错误。

1.1 科级水平

如表1所示，通过对样品自编号为Z5376和Z5377的7条DNA条形码序列进行BLAST比对，发现样品编号Z5376原标本鉴定信息为禾本科雀稗（*Paspalum thunbergii*），但Z5376的3条DNA条形码序列（*rbc*L、*mat*K和*trn*H-*psb*A）经BLAST搜索与香蒲科香蒲属（*Typha*）植物相似性最高；而来源于Z5377的4条DNA条形码序列原鉴定信息为长苞香蒲（*Typha angustata*），经BLAST搜索则与禾本科雀稗属（*Paspalum*）序

列信息相似。为了进一步对 BLAST 结果验证，利用建树法对这两个属目前已获得的 DNA 条形码序列构建 NJ 树，也发现样品 Z5376 和 Z5377 与来源于同属的其它样品间遗传距离较大，枝长较长，确定这两份样品为错误鉴定（图 1）。

1.2　属级水平

如图 2 所示，基于数据库中已有条形码数据信息，对马鞭草科 *mat*K 数据比对构建 NJ 树，结果显示紫珠属（*Callicarpa*）、莸属（*Caryopteris*）、马缨丹属（*Lantana*）和马鞭草属（*Verbena*）构成了较好的单系支，牡荆属（*Vitex*）为并系，而赪桐属（*Clerodendrum*）编号为 D0705 的样品被鉴定为三对节（*Clerodendrum serratum*），其遗传距离与牡荆属植物更相似；同样，根据 BLAST 结果，本样品与牡荆属物种的序列相似度最高，也支持该样品应为牡荆属植物（表 1）。通过对该样品标本的查证，D0705 样品标本鉴定有误。

1.3　种级水平

如图 3 所示，对接骨木属（*Sambacus*）已有 *mat*K 数据进行序列比对，构建 NJ 树，结果发现血满草（*S. adnata*）的 9 份样品聚类在一起；接骨草（*S. javanica*）的 16 份样品除 D3577 和 Z0300 外聚为一支，且序列完全相同，样品 D3577 单独成为一支，而样品 Z0300 则与血满草聚类为一支，而且序列相同。BLAST 结果支持这两份样品均为接骨木属植物（表 1）。通过对样品标本查证，表明样品 Z0300 确实为血满草，样品 D3577 物种鉴定错误。

2　对未鉴定标本的辅助鉴定

一些标本由于采集性状不完整，缺乏物种关键鉴定特征，如花、果等，给传统分类学的鉴定带来困难。DNA 条形码有诸多优点，如物种鉴定不受个体形态特征的限制（如：形态性状特征不显著的一小块或一小片残缺材料；保存已久的腊叶标本、炮制过的民间药材、DNA 降解材料等）；鉴定结果准确，具有可重复性。但只有构建统一完整的 DNA 条形码参考数据库，才能快速实现对已知物种的准确鉴定以及新种和隐存种的发现。

借助已有较好 DNA 条形码研究类群的数据库，我们尝试通过 DNA 条形码技术，将这些材料进行准确的物种鉴定。例如，Liu 等（2011）利用 4 个叶绿体片段（*rbc*L、*mat*K、*trn*L-F、*trn*H-*psb*A）和 1 个核基因片段（ITS），将分布于欧亚地区的红豆杉属（*Taxus*）植物鉴定为 11 个种，即欧洲红豆杉（*T. baccata*）、密叶红豆杉（*T. fuana*）、东北红豆杉（*T. cuspidata*）、苏门答腊红豆杉（*T. sumatrana*）、喜马拉雅红豆杉（*T. wallichiana* s. s.）、红豆杉（*T. chinensis*）、南方红豆杉（*T. mairei*），和 4 个新分类群或隐存种，

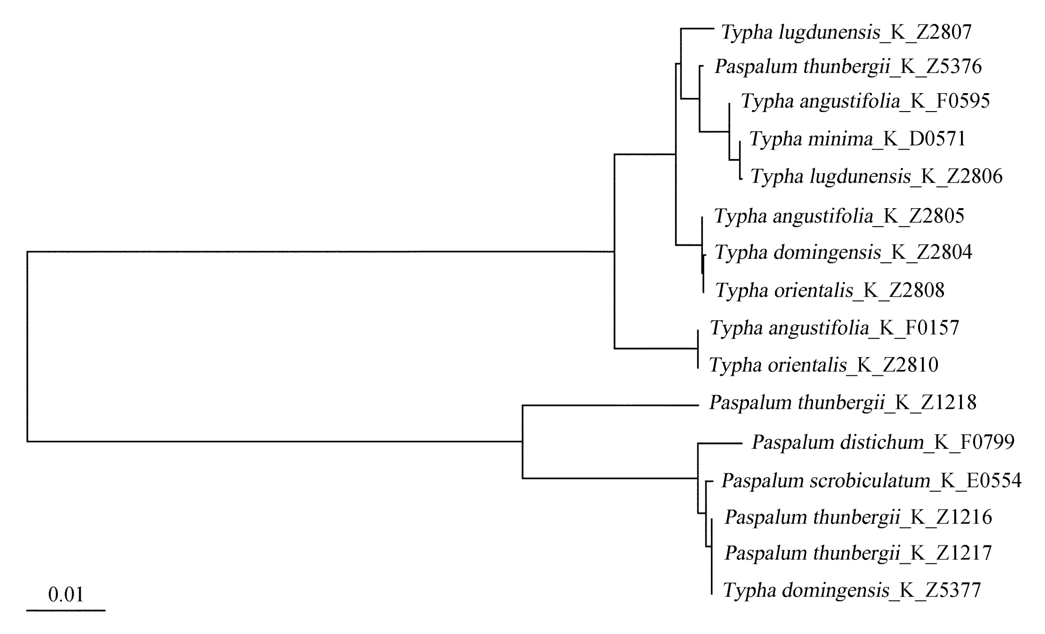


图 1　基于 *mat*K 序列用 *p* 距离构建的邻接树

Fig. 1　Unrooted neighbor-joining (NJ) tree based on *mat*K data with *p*-distance

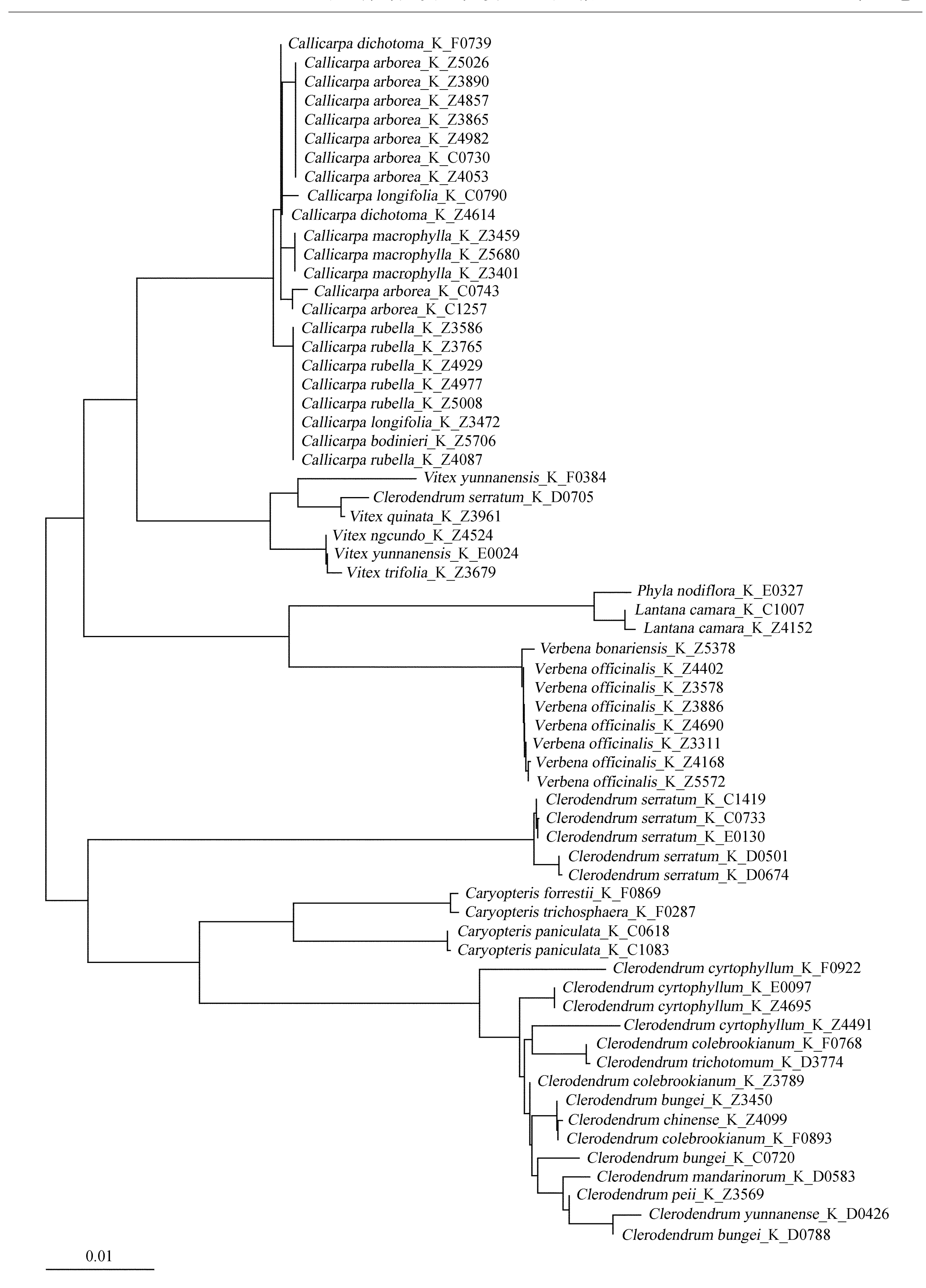


图2 基于 *mat*K 序列用 *p* 距离构建的邻接树

Fig. 2 Unrooted neighbor-joining (NJ) tree based on *mat*K data with *p*-distance

表 1 样品编号 Z5376、Z5377、D0705、Z0300 和 D3577 BLAST 分析结果

Table 1 BLAST results of samples of Z5376、Z5377、D0705、Z0300 and D3577

Query_ID (sample ID_barcode)	Query_species	Blast_species (GI number_species name)	Query Coverage	Identity
Z5376_*rbc*L	*Paspalum thunbergii*	218175479_*Typha domingensis*	718	99.86
Z5376_*mat*K	*Paspalum thunbergii*	90967927_*Typha angustifolia*	861	99.65
Z5376_*trn*H-*psb*A	*Paspalum thunbergii*	166244620_*Typha angustata*	703	98.86
Z5377_*rbc*L	*Typha angustata*	341872831_*Paspalum thunbergii*	720	99.72
Z5377_*mat*K	*Typha angustata*	440576753_*Paspalum urvillei*	813	99.26
Z5377_*trn*H-*psb*A	*Typha angustata*	341939337_*Paspalum thunbergii*	641	99.84
Z5377_ITS	*Typha angustata*	341872750_*Paspalum thunbergii*	599	99.67
D0705_*rbc*L	*Clerodendrum serratum*	290586059_*Vitex negundo*	678	99.41
D0705_*mat*K	*Clerodendrum serratum*	24430461_*Vitex quinata*	755	99.34
D0705_*trn*H-*psb*A	*Clerodendrum serratum*	331704877_*Vitex quinata*	435	99.31
Z0300_*rbc*L	*Sambucus javanica*	306482292_*Sambucus* sp.	643	100
Z0300_*mat*K	*Sambucus javanica*	379042244_*Sambucus ebulus*	775	99.87
Z0300_*trn*H-*psb*A	*Sambucus javanica*	343196454_*Sambucus chinensis*	392	99.74
Z0300_ITS	*Sambucus javanica*	343406909_*Sambucus chinensis*	611	99.84
D3577_*rbc*L	*Sambucus javanica*	306482292_*Sambucus* sp.	718	99.44
D3577_*mat*K	*Sambucus javanica*	379042244_*Sambucus ebulus*	822	99.76
D3577_*trn*H-*psb*A	*Sambucus javanica*	343196452_*Sambucus chinensis*	406	97.78
D3577_ITS	*Sambucus javanica*	343406916_*Sambucus chinensis*	610	99.18

即 Hengduan type（= *T. florinii* Spjut），Qinling type，Tonkin type（= *T. colcicola*）和 Emei type（Moeller 等，个人通讯）（图 4）。我们建立的数据库中，红豆杉属材料共 7 份，即 C0293、D2198、D2199、D2200、D2205、E0960 和 Z2719，基于形态学鉴定均为 *Taxus* sp.。基于这 7 份样品已有的条形码信息（ITS、*rbc*L 和 *trn*H-*psb*A），从 NCBI 下载获得红豆杉属 47 个个体的 ITS、*rbc*L 和 *trn*H-*psb*A 序列，构建联合矩阵，NJ 树结果如图 5 所示。尽管仅使用了 3 个 DNA 条形码序列，红豆杉属的 11 个种都构成了单系支。数据库中测试样品 D2205 与密叶红豆杉聚在一起，C0293、D2199、D2200 和 Z2719 与南方红豆杉聚为一支，D2198 和 E0960 则与 Tonkin type 遗传距离更接近。由此，我们认为 D2205 应为密叶红豆杉，C0293、D2199、D2200 和 Z2719 为南方红豆杉，E0960 和 D2198 为灰岩红豆杉 Tonkin type。

3 植物 DNA 条形码应用于物种鉴定的几点思考

DNA 条形码鉴定必要的前提是基于分类群变异合适的标准条形码序列，构建一个完备的物种 DNA 条形码参考数据库。构建植物 DNA 条形码参考数据库的过程中，有下列三个方面值得关注。

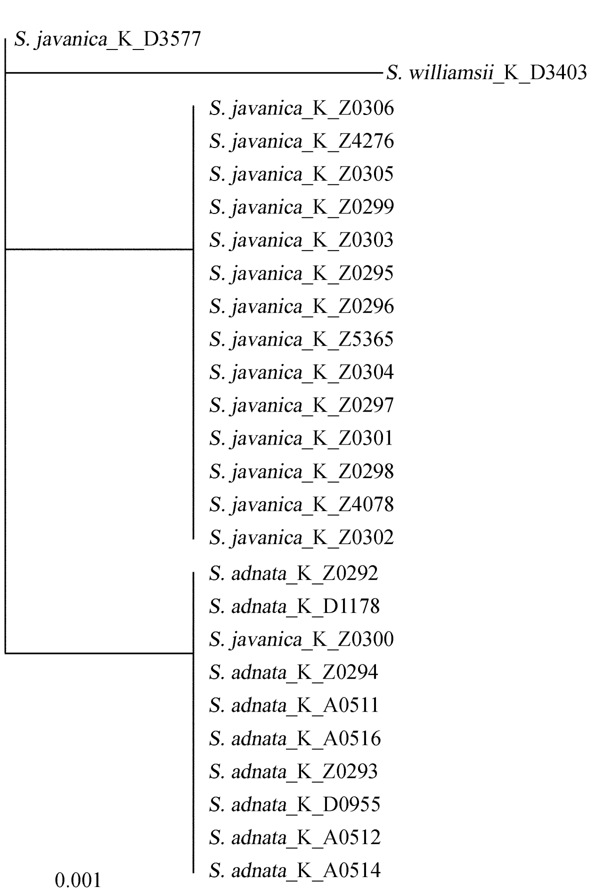


图 3 基于 *mat*K 序列用 *p* 距离构建的邻接树

Fig. 3 Unrooted neighbor-joining (NJ) tree based on of *mat*K data with *p*-distance

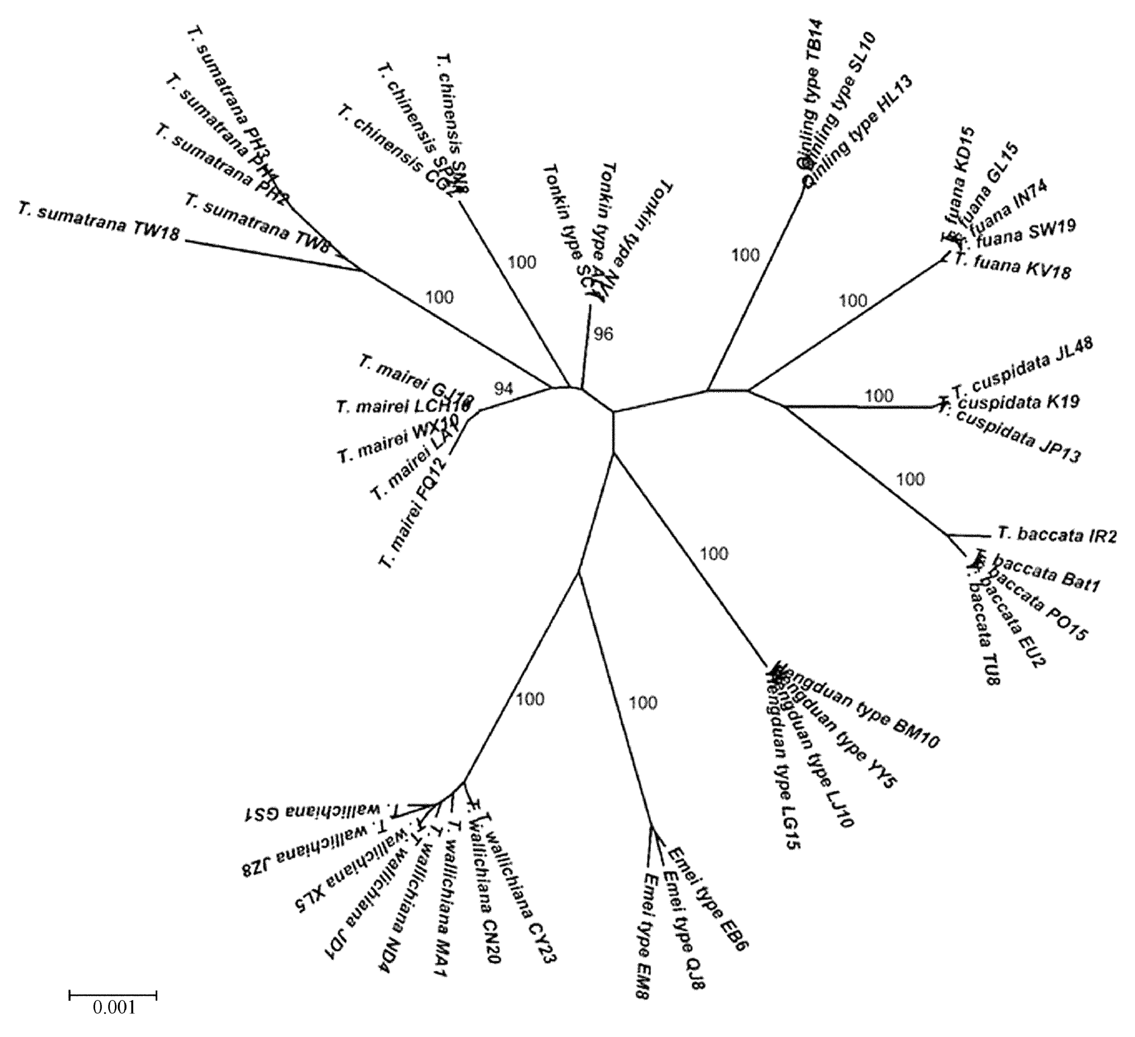


图4　基于 *rbc*L、*mat*K、*trn*H-*psb*A、*trn*L-F 和 ITS 联合矩阵用 *p* 距离构建的邻接树（图片引自 Liu 等，2011）

Fig. 4　Unrooted neighbor-joining（NJ）tree based on *rbcL*、*matK*、*trnH-psbA*、*trnL-F* and ITS data with *p*-distance（Figure cited from Liu *et al*. 2011）

3.1　物种鉴定准确

DNA 条形码数据库的构建，第一步是物种材料的采集。形态分类学研究是正确取样的基础（Meyer 和 Paulay，2005；DeSalle，2006）。同时为了保证物种的准确性，取样还应涵盖物种的整个分布区，充分考虑居群的概念，准确和较密集取样，尽可能最大限度地代表物种遗传多样性（高连明等，2012）。

3.2　物种与 DNA 条形码对应关系准确

首先，数据收集和管理应标准化和信息化。DNA 条形码数据库构建过程中涉及样品采集、DNA 提取、PCR 扩增、测序与拼接和序列分析。每一步骤都应制定和遵守相应的标准规范，如物种命名、凭证标本的存放、图片及地理信息、条形码序列编号和管理、数据分析等。同时，与计算机网络技术相结合，样品入库出库、数据产生和提取的过程，引入信息化的数据管理。一方面便于数据的共享、查询和实时更新，另一方面避免人为因素导致的样品混淆。

其次，DNA 条形码与实物库相关联，利用 DNA 条形码分析手段对凭证标本重新鉴定，检验形态分类结果。DNA 条形码序列必须与实物库，即标本库、植物分子材料库和植物总 DNA 库相关联。通过 DNA 条形码分析手段，将分析结果与原形态鉴定结果相比较，查询凭证标本的采集信息（地理分布、生境、海拔）、标本图片等，对凭证标本再鉴定。将形态分类与 DNA 条形码研究紧密结合，保证 DNA 条形码序列与物种的正确对应关系。

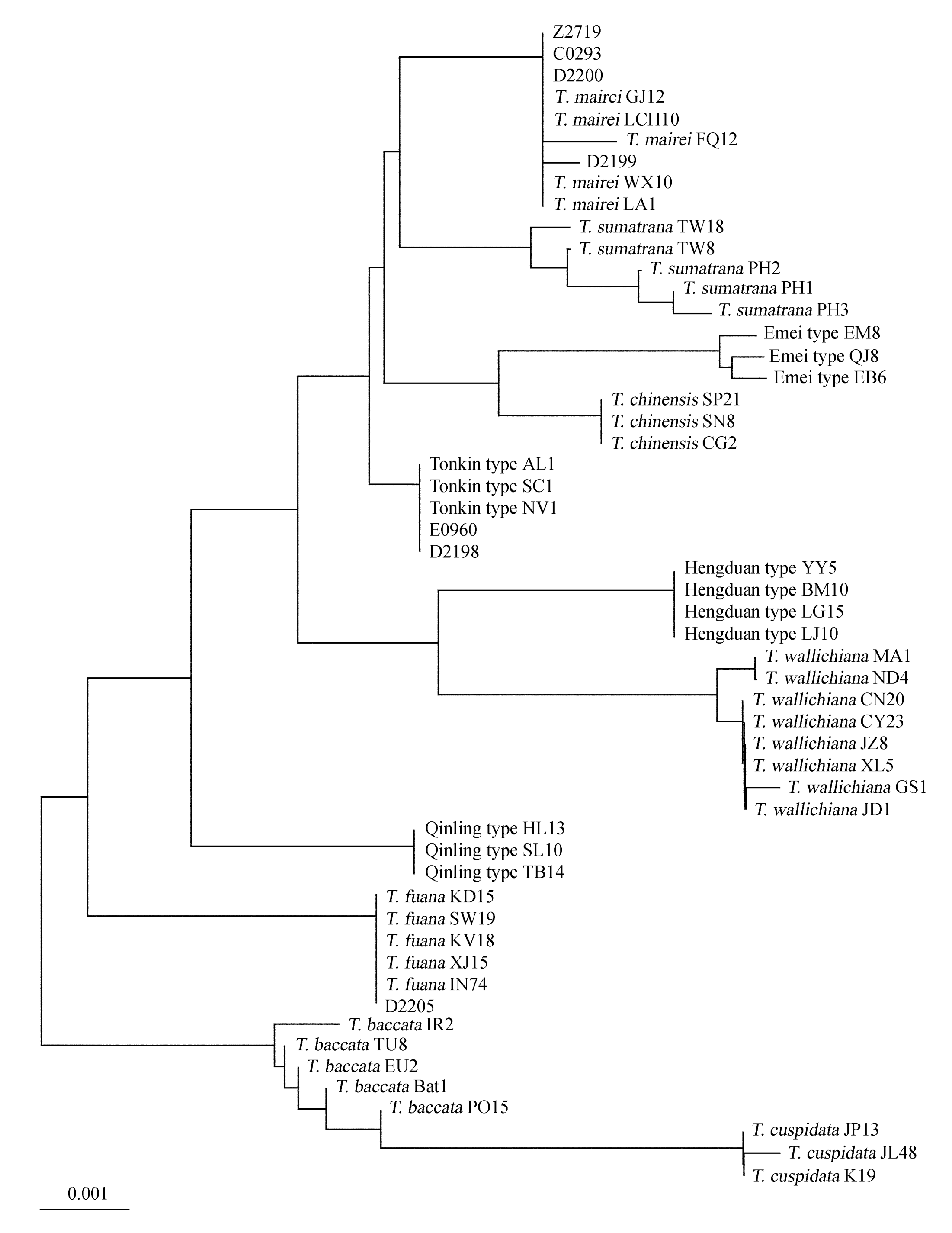


图 5　基于 ITS、*rbc*L 和 *trn*H-*psb*A 联合矩阵用 *p* 距离构建的邻接树

Fig. 5　Unrooted neighbor-joining (NJ) tree based on ITS、*rbc*L and *trn*H-*psb*A data with *p*-distance

3.3　演化关系复杂的类群需要多种方法验证

植物的演化较为复杂，杂交或者基因渐渗、近期辐射分化等现象较为普遍，均可能导致物种的界限模糊不清，现有的 DNA 条形码难以有效解决等。这些困难类群的鉴定，还需要开展有针对性的探索研究，如筛选其它合适候选条码或分子标记等，从而为构建完整的条形码数据库做好充分的准备。

致谢　感谢中国科学院昆明植物研究所王红研究员在本文撰写过程中提出的宝贵意见；分子生物学实验中心袁文斌助理工程师、林春艳助理工程师和张金萍助理工程师在 DNA 条形码数据获取与拼接中的帮助。

〔参 考 文 献〕

Agnarsson I, Kuntner M, 2007. Taxonomy in a changing world: seeking solutions for a science in crisis [J]. *Systematic Biology*, **56**: 531—539

Blaxter M, 2003. Molecular systematics-counting angels with DNA [J]. *Nature*, **421**: 122—124

CBOL Plant Working Group, 2009. A DNA barcode for land plants [J]. *Proceedings of the National Academy of Sciences of the United States of America*, **106**: 12794—12797

Chase MW, Fay F, 2009. Barcoding of plants and fungi [J]. *Science*, **325**: 682—683

Che J, Huang DW, Li DZ *et al*., 2010. DNA barcoding and the international barcode of life project in China [J]. *Bulletin of Chinese Academy of Sciences* (中国科学院院刊), **24**: 257—560

Dayrat B, 2005. Towards integrative taxonomy [J]. *Biological Journal of the Linnean Society*, **85**: 407—415

DeSalle R, 2006. Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff [J]. *Conservation Biology*, **20**: 1545—1457

DeSalle R, Egan MG, Siddall M, 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding [J]. *Philosophical Transactions of the Royal Society of London-Series B: Biological Sciences*, **360**: 1905—1916

Fang W (方伟), Liu ED (刘恩德), 2012. The development of classical plant taxonomy and iFlora [J]. *Plant Divesity and Resources* (植物分类与资源学报), **34** (6): 525—531

Fitzhugh K, 2006. DNA barcoding: an instance of technology-driven science? [J]. *Bioscience*, **56** (6): 462—463

Gao LM (高连明), Liu J (刘杰), Cai J (蔡杰) *et al*., 2012. A synopsis of technical notes on the standards for plant DNA barcoding [J]. *Plant Divesity and Resources* (植物分类与资源学报), **34** (6): 592—606

Godfray HCJ, 2002. Towards taxonomy's 'glorious revolution' [J]. *Nature*, **420**: 461

Haase M, Wilke T, Mildner P, 2007. Identifying species of *Bythinella* (Caenogastropoda: Rissooidea): a plea for an integrative approach [J]. *Zootaxa*, **1563**: 1—16

Hajibabaei M, Singer GAC, Hebert PDN *et al*., 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics [J]. *Trends in Genetics*, **23**: 167—172

Hawksworth DL, 1995. Challenges in mycology [J]. *Mycological Research*, **99**: 127—128

Hebert PDN, Cywinska A, Ball SL *et al*., 2003. Biological identifications through DNA barcodes [J]. *Proceedings of the Royal Society of London-Series B, Biological Sciences*, **270**: 313—321

Hebert PDN, Gregory TR, 2005. The promise of DNA barcoding for taxonomy [J]. *Systematic Biology*, **54** (5): 852—859

Hollingsworth PM, Graham SW, Little DP, 2011. Choosing and using a plant DNA Barcode [J]. *PLoS ONE*, **6** (5): e19254

Kristiansen KA, Cilieborg M, Drabkova L *et al*., 2005. DNA taxonomy-the riddle of *Oxychloe* (Juncaceae) [J]. *Systematic Botany*, **30** (2): 284—289

Li DZ, Gao LM, Li HT *et al*., 2011a. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants [J]. *Proceedings of the National Academy of Sciences of the United States of America*, **108**: 19641—19646

Li DZ, Liu JQ, Chen ZD *et al*., 2011b. Plant DNA barcoding in China [J]. *Journal of Systematics and Evolution*, **49**: 165—168

Liu J, Michael Möller, Gao LM *et al*., 2011. DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species [J]. *Molecular Ecology Resources*, **11** (1): 89—100

Mallet J, Willmott K, 2003. Taxonomy: renaissance or tower of Babel? [J]. *Trends in Ecology and Evolution*, **18**: 57—59

Matz MV, Nielsen R, 2005. A likelihood ratio rest test for species membership based on DNA sequence data [J]. *Philosophical Transactions of the Royal Society of London-Series B: Biological Sciences*, **360**: 1969—1974

Meier R, Shiyang K, Vaidya G *et al*., 2006. DNA barcoding and taxonomy in diptera: a tale of high intraspecific variability and low identification success [J]. *Systematic Biology*, **55** (5): 715—728

Meyer CP, Paulay G, 2005. DNA barcoding: error rates based on comprehensive sampling [J]. *PLoS Biology*, **3** (12): e422

Mora C, Tittensor DP, Adl S *et al*., 2011. How many species are there on earth and in the ocean? [J]. *PLoS Biology*, **9** (8): e1001127

Nielsen R, Matz M, 2006. Statistical approaches for DNA barcoding [J]. *Systematic Biology*, **55** (1): 162—169

Ren BQ (任保青), Chen ZD (陈之端), 2010. DNA barcoding plant life [J]. *Chinese Bulletin of Botany* (植物学报), **45** (1): 1—12

Schindel DE, Miller SE, 2005. DNA barcoding a useful tool for taxonomists [J]. *Nature*, **435**: 17

Tautz D, Arctander P, Minelli A *et al*., 2003. A plea for DNA taxonomy [J]. *Trends in Ecology & Evolution*, **18** (2): 70—74

Vogler AP, Monaghan MT, 2007. Recent advances in DNA taxonomy [J]. *Journal of Zoological Systematics and Evolutionary Research*, **45**: 1—10

Will KW, Mishler BD, Wheeler QD, 2005. The perils of DNA barcoding and the need for integrative taxonomy [J]. *Systematic Biology*, **54** (5): 844—851